DOCUMENT RESUME

ED 263 180 TM 850 670

AUTHOR	Smith, Nick L.; St. John, Mark
TITLE	Overview of Alternative Evaluation Methods. Paper and Report Series, No. 107. Interim Draft.
INSTITUTION	Northwest Regional Educational Lab., Portland, OR. Research on Evaluation Program.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
REPORT NO	NWREL-REP-107
PUB DATE	Mar 85
CONTRACT	400-80-0105
NOTE	37p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Elementary Secondary Education; Evaluation Criteria; *Evaluation Methods; *Evaluation Needs; Evaluators; *Guidelines; *Measurement Techniques; Program Evaluation; *Research Methodology; Use Studies
IDENTIFIERS	*Northwest Regional Educational Laboratory

ABSTRACT

The Research on Evaluation Program's National Advisory Panel recommended that a means of matching evaluation methods with evaluation purposes be developed and that it include a guide to major Program publications. Practitioners in the field have also requested support on how to approach the evaluation design process to permit the incorporation of alternative evaluation approaches. The present report has been designed to meet these requests. Provided is an introductory overview to the range of evaluation methods currently available, as well as guidelines to assist practitioners in making their own evaluative reviews and selections of appropriate methods. To promote the widest possible distribution of this material, the two sections of this report have been designed as field guides. After final revisions, they will be released as part of the Program's "Evaluation Guides" series. The first section discusses procedure for selecting from among alternative evaluation methods and includes numerous table displays to aid in that process. The second section presents a strategy for designing evaluation studies so that the evaluation methodology most appropriate to the problem at hand may be selected. (PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. M. Rogers

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Interim Draft

Do not cite or quote without authors' permission.

Comments and suggestions are welcomed.

No. 107 OVERVIEW OF ALTERNATIVE EVALUATION METHODS

NICK L. SMITH
MARK ST. JOHN

March 1985

Nick L. Smith, Director
Research on Evaluation Program
Northwest Regional Educational Laboratory
300 S.W. Sixth Avenue, Portland, Oregon 97204

Published by the Northwest Regional Educational Laboratory, a private nonprofit corporation. The work upon which this publication is based was performed pursuant to Contract No. 400-80-0105 of the National Institute of Education. It does not, however, necessarily reflect the views of that agency.

The information presented in this publication does not necessarily reflect the opinions of the Northwest Regional Educational Laboratory and no endorsement should be inferred.

PREFACE

The Research on Evaluation Program is a Northwest Regional Educational Laboratory project of research, development, testing, and training designed to create new evaluation methodologies for use in education. This document is one of a series of papers and reports produced by program staff, visiting scholars, adjunct scholars, and project collaborators--all members of a cooperative network of colleagues working on the development of new methodologies.

Over the past several years, the Research on Evaluation Program has conducted research and development on a variety of alternative evaluation methods. This report contains an introductory overview of those methods, written to acquaint the reader with the wide range of evaluation approaches currently available. After final revisions, the majority of this report will be re-issued as two guides for (a) selecting new evaluation methods, and (b) designing studies using them.

Nick L. Smith, Editor
Paper and Report Series

CONTENTS

	Page
INTRODUCTION	1
References	6
SELECTING ALTERNATIVE METHODS	7
The Need for a Repertoire of Methods	9
Guidelines for Selecting Methods	10
Summary	14
References	21
DESIGN STRATEGY FOR INCORPORATING NEW METHODS	25
The Design Problem	27
The Architect	28
The Evaluator	31
Summary	33
Reference	35

INTRODUCTION

Over the past several years, the Research on Evaluation Program has conducted research and development on a wide variety of alternative evaluation approaches. Table 1 lists 52 tools within 17 general method areas which the Program has investigated.

We had originally intended to conduct an evaluative review of each method in terms of its potential for improving evaluation practice in state departments of education and local school districts. This task has proven unmanageable for a number of reasons. First, our studies of evaluation practice (Smith, 1982; Gray, Caulley, and Smith, 1982; Smith, J. K, 1984; Smith, N. L., 1984; Smith and Smith, in press) have repeatedly emphasized the diversity of settings within which educational evaluation takes place. What is an acceptable and effective method at one site might be totally unacceptable at another. Second, to focus only on ROEP work on a given method would be to ignore the important work being done throughout the U. S. and abroad to develop improved evaluation methods. However, to include the significant other work being done on these methods is difficult because of the rapidity with which progress is currently being made in method improvement, not to mention that the resources required would far exceed those available for completion of the current task.

More importantly, we have increasingly received requests from practitioners and researchers using our materials, for assistance in sorting through the variety of methods we have identified. The Program's National Advisory Panel has recommended that we develop a means of matching evaluation methods with evaluation purposes and include a guide to our major Program publications. Practitioners in the field have also requested support on how to approach the evaluation design process in such a way as to permit the incorporation of alternative evaluation approaches.

The present report has been designed to meet these requests. We have provided here an introductory overview to the range of evaluation methods currently available and have provided guidelines to assist practitioners in making their own evaluative reviews and selections of appropriate methods. In order to promote the widest possible distribution of this material, the following two sections of this report have been designed as field guides. After final revisions, they will be released as part of the Program's Evaluation Guides series.

The following section discusses procedures for selecting from among alternative evaluation methods and includes numerous table displays to aid in that process. The final section presents a strategy for designing evaluation studies so that the evaluation methodology most appropriate to the problem at hand may be selected.

Table 1

The Range of Methods and Tools Investigated by the Research on Evaluation Program

General Methods	Specific Tools
Art/Design	Composition Graphics
Case Study	Interviewing Field observations Qualitative analysis Aggregation techniques Vignettes
Causal Research	Experimental designs Quasi-experimental designs Path analysis
Cost Analysis	Feasibility analysis Utility analysis Benefit analysis Effectiveness analysis
Criticism	Thematic matrix analysis Connoisseurship Composing techniques
Document Analysis	Tracking Legislative history Content analysis
Exploratory Data Analysis	Stem and leaf displays Box plots Functional transformation of data
Geographic Methods	Geocode analysis Trend surface analysis Social area analysis Maps
Hearings	Committee hearings Panel reviews
Investigative Journalism	Quick study Interviewing Tracking Shuffling, circling Files and profiles

(cont.)

Table 1 cont.

General Methods	Specific Tools
Operations Research	Assignment models Transportation models Queuing theory
Philosophical Analysis	Concept analysis
Photography	Photo interviewing Sampling techniques
Product Evaluation	Needs assessment Critical competitor Side effects checks
Service Delivery Assessment	Qualitative methods Debriefings Briefings
Storytelling	Oral histories Narrative techniques Stream of consciousness
Survey Research	Questionnaires Surveys Interviews

REFERENCES

Gray, P. J., Caulley, D. N., & Smith, N. L. (1982). A study in contrasts: Effects of the Education Consolidation and Improvement Act of 1981 on SEA and LEA evaluation (Research on Evaluation Program Paper and Report Series No. 79). Portland, OR: Northwest Regional Educational Laboratory.

Smith, J. K. (1984). Cost analysis at the local level: Applications and attitudes (Research on Evaluation Program Paper and Report Series No. 103). Portland, OR: Northwest Regional Educational Laboratory.

Smith, N. L. (1982). The context of evaluation practice in state departments of education. Educational Evaluation and Policy Analysis, 4(3), 373-386.

Smith, N. L. (1984). Evaluation units in state departments of education: A five-year update. Evaluation News, 5(3), 37-44.

Smith, N. L. & Smith, J. K (in press). State level evaluation uses of cost analysis: A national descriptive survey. In J. Catterall (Ed.), Economic analysis of public programs. San Francisco: Jossey-Bass.

SELECTING ALTERNATIVE METHODS

THE NEED FOR A REPERTOIRE OF METHODS

When three blindmen in India set out to discover the true nature of the elephant, they all found it to be quite different--one said it was like a wall; another found it like a snake; a third was sure that elephants were like trees. By looking in one way, in one place, each thought he had understood the whole.

Evaluators have the same tendency. By identifying themselves with an approach or a method, they limit their senses and restrict their view. And since what they see depends on how they look, they are able (however accurately) to learn about only one aspect of the program. Thus, when evaluators adopt a one-dimensional approach, they end up in a position where they are able to claim "the program is like this. . ." If, on the other hand, evaluators have at their command a wide repertoire of approaches and methods, and if they know how to combine these methods appropriately and flexibly, then, at the very least, they will be able to say, "The program is like this, and this, and this . . ."

But you may argue, "Of course it is better to have more information, and therefore more views of a program, but resources are limited. We have to choose; we need to decide which methods really work, and which are really feasible and useful in the situations we face."

This is true, but it is not as absolutely true as we usually believe. That is, we do not always need to choose a single method for all situations. To understand this, consider, as an analogy, the way that a carpenter works.

A carpenter does not choose whether a hammer or saw is more useful--he knows they are both tools of his trade, and they are equally and exclusively useful, depending upon the situation. Carpenters do not find a need to debate the relative merits of the hammer and saw. Nor do carpenters have a preference for one tool over the other; they don't try to use a hammer when a saw is required.

Different evaluation methods (experiments, case studies, surveys . . .) can be seen to be like the carpenter's hammer and saw--they are the evaluator's tools. Evaluators would do well to think of themselves as artists and craftsmen, and take pride in learning to use a wide range of tools skillfully. As carpenters

do, evaluators can learn to eye a situation and know instinctively which tools will work. As their skill increases, evaluators can learn how to extend the range of use of their tools, and even how to combine their uses in innovative ways to accomplish more difficult tasks.

Carpenters' tools have evolved over centuries of continuous use; consequently, modern carpentry tools are simple, elegant, and ideally suited for their uses. By comparison, evaluators have had a short history of tool development.

Fortunately, other disciplines have developed and refined methods for evaluation in their own fields. For example, wine tasters, film critics, accreditation agencies, investigative journalists, test drivers, and senate committees have all developed approaches that are potentially useful for the educational evaluator. For several years, the Research on Evaluation Program has been collecting and adapting these kinds of methods for use in the evaluation of educational, social, or health programs. A sizeable literature describing the evaluation methods and techniques of a wide range of disciplines now exists. For the evaluator who is willing to experiment and learn new skills, these methods can become very useful tools.

GUIDELINES FOR SELECTING METHODS

The development of many new and varied methods is perhaps a mixed blessing for the evaluator. For, while it empowers the evaluator, it also makes the task of deciding what to do more difficult. When should the evaluator experiment with a new method? and which method? What is the basis on which choices are to be made?

To answer these questions, it is useful to remember that the choice of methods is part of the overall evaluation design process which can be thought to consist of:

1. analyzing the problem context;
2. asking a few general questions;
3. selecting the methods (strategies) to use;
4. selecting the specific techniques (tactics) to use.

Analyzing the Context

Like any other design process, an evaluation begins with an analysis of the problem context--that is, with an investigation of the program as it is, and with an attempt to learn what is important in the setting. In a typical case, an evaluator faced with complexity may become lost in all the factors of problem context, as shown in Figure 1.

As the evaluator investigates each of these factors and begins to "get the lay of the land," a working plan (a design) begins to form in his or her mind. Moving from the general to the specific, the evaluator begins to form some general and guiding questions in his or her mind. These questions provide a focus and purpose to the evaluation. They prioritize what is important to study. Examples of such general questions include:

1. How can we best understand what is happening in this program?

2. How could this program be made to work better?

3. What are the outcomes of this program?

4. What important variations are there in the program's activities or effects?

5. How worthwhile is the program?

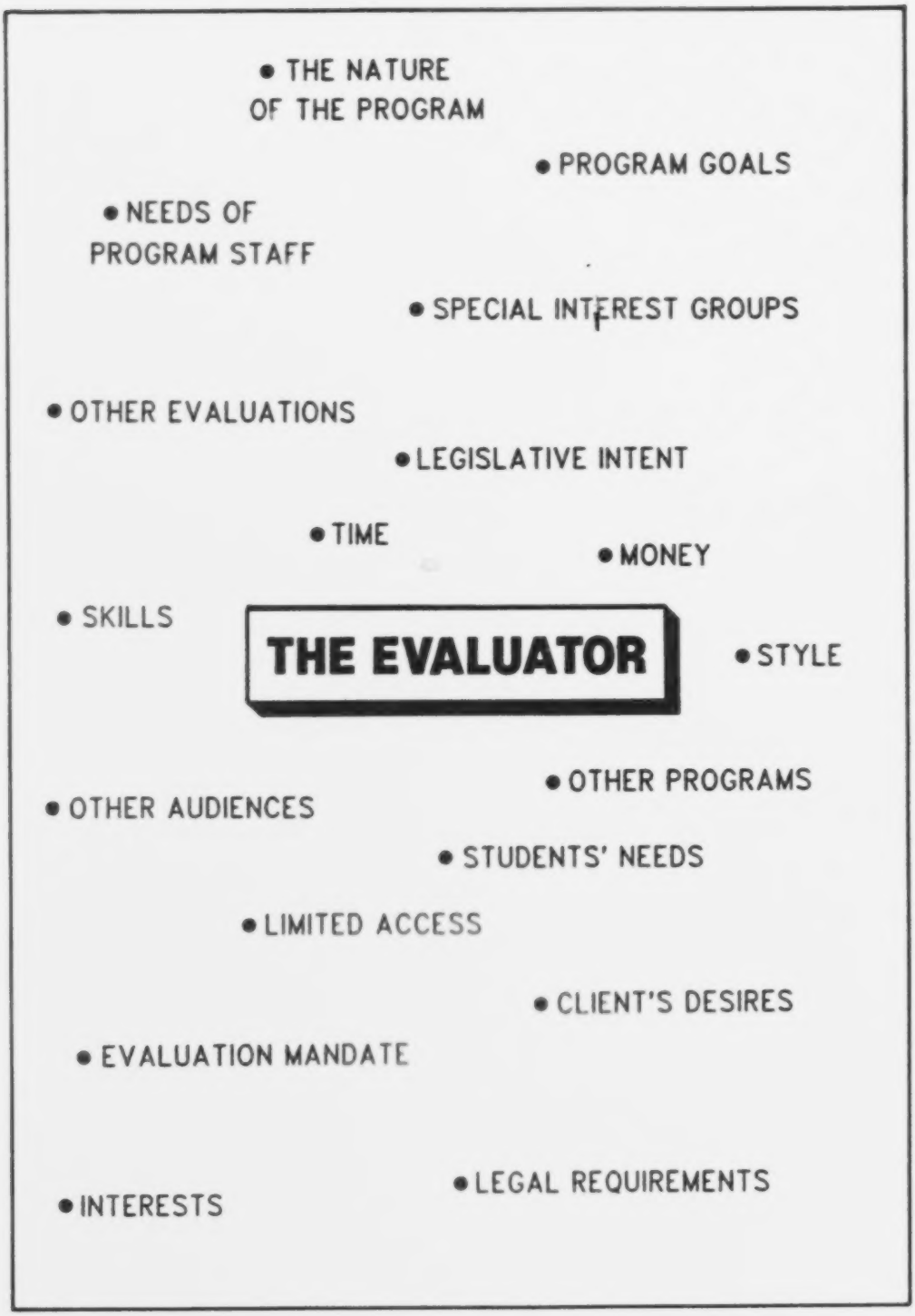
Example:

An evaluator in a state evaluation agency is given the task of evaluating a fourth grade reading program. The predominant factors in the problem context are the requirements set by the legislature for evaluation and the political pressure of parental groups for basic skills improvement. Thus, the evaluator in this case has little leeway--the general question that must be addressed is something like this: how much have the fourth grade students improved their reading skills over the year?

Example:

A museum has been given funding to train middle school science teachers. The evaluator has been called in early in the program to assist in any way possible. Here the evaluation problem context is more loosely defined, and the predominant general questions that emerge are: (1) What is happening in the program? (2) Who are these teachers? and What are their actual needs? and (3) How can the practices of the program be improved?

Figure 1


• THE NATURE OF THE PROGRAM
• PROGRAM GOALS
• NEEDS OF PROGRAM STAFF
• SPECIAL INTEREST GROUPS
• OTHER EVALUATIONS
• LEGISLATIVE INTENT
• TIME
• MONEY
• SKILLS
THE EVALUATOR
• STYLE
• OTHER PROGRAMS
• OTHER AUDIENCES
• STUDENTS' NEEDS
• LIMITED ACCESS
• CLIENT'S DESIRES
• EVALUATION MANDATE
• LEGAL REQUIREMENTS
• INTERESTS

Selecting Methods

When the evaluator begins to pose general questions, he or she is implicitly beginning to select responsive and appropriate methods. That is, certain methods are better than others for answering the relevant questions of the evaluation. In the previous example of the evaluator assessing the improvement of the reading skills of the fourth graders, experimental methods and achievement tests are obviously called for. The evaluator working with the museum would more likely use case study or journalistic methods.

The choice of methods also suggests a stance or role that the evaluator may adopt. Usually, almost unconsciously, evaluators adopt the role of the experimenter and tester. In using other methods, evaluators may bring a different approach with different perspectives and values to the evaluation. The evaluator can adopt not only the techniques, but also the perspective of the discipline from which the method is derived. Thus, in using journalistic methods, the evaluator begins to think and act like a journalist; in using methods of operations research, the evaluator takes on the role of the efficiency expert.

Techniques

If methods reflect the strategies of evaluators, techniques reflect the tactics. Methods come from the parent discipline, and they are complete with rationale, perspective, and techniques. The techniques are more specific and can be viewed as tools--the evaluator's tools. Techniques are adaptive and flexible in their use. They are used for gathering and analyzing data, for organizing findings, and for presenting results. Interviews, t-tests, surveys, research briefs, and thematic analysis are all examples of techniques used in evaluation.

Putting it all Together

As the evaluator gains an understanding of the evaluation context, general questions begin to emerge. These questions in turn suggest methods and approaches with which to structure the evaluation. Within the context of the methods, specific techniques are used by the evaluator. The chart below shows a few examples of how these elements go together to form a rough working plan.

General Question(s)	Method(s)	Technique(s)	Evaluator's Stance
What are the main issues for the program participants?	Case Study	Interviews	Anthropologist
How could the program be run more efficiently?	Modeling	Queueing theory	Operations Researcher
Is this method more effective?	Pre-post Control	Analysis of variance	Experimental Designer
What is it like to be in this program?	Story-telling	Stream of consciousness	Storyteller

This movement from general questions to methods and techniques is illustrated more fully in Table 1. Part A of Table 1 lists five general questions and associated methods that cover a wide spectrum of evaluation purposes. Part B of Table 1 suggests questions and methods that can help the evaluator in the design process itself.

Table 2 lists more specific evaluation techniques (tools), describing the discipline from which they arise and the purpose for which they are most suited.

SUMMARY

Because the evaluation process can be viewed as a problem in design, the following conclusions can be made about the planning and implementation of an evaluation:

- The evaluation problem is undefined; there is no one unique solution.
- Successful evaluation design cannot be prescribed; at best, a few heuristics may prove useful.
- The power and flexibility of the evaluator is increased as his or her repertoire of methods and skills increases.

Thus, the evaluator's search for a best method or even a best way to choose among methods is futile. Rather, the evaluator is left in a less well defined but freer world. To operate successfully in this world, what the evaluator needs is a knowledge of many different kinds of methods, an understanding of their different purposes, and an appreciation of their limitations--and then, in addition, a little courage and imagination . . .

Table 1

A. Selecting Methods

I: How Can We Best Understand What is Happening in the program?
What is the nature or character of the program?
What are the conditions and activities like?
What are the central issues, themes, conflicts, trade-offs?
What seems important?

Useful Methods	Relevant Tools	General Purpose	References*
Investigative Journalism	Interviewing Tracking Shuffling, circling Files and profiles	To discover that which is important, but not immediately apparent	8, 15
Case Study	Interviewing Field observations Qualitative analysis Aggregation techniques	To gain insight into a program by understanding the many facets of the actual (not ideal) functioning of the program	
Storytelling	Oral histories Narrative techniques Stream of consciousness	To convey humanness and complexity of a program; to create images and establish mental connections	22, 26
Criticism	Thematic matrix analysis Connoisseurship Composing techniques	To illuminate forms; to demystify; to enhance sensibilities To provide an artful representation of reality	5, 6
Photography	Photo interviewing Sampling techniques	To capture, and portray images of reality; to illustrate themes or issues; to deepen insights with visual images	25
Exploratory Data Analysis	Stem and leaf displays Box plots Functional transformation of data	To discover relationships not immediately apparent in accumulated data; to select appropriate analytical methods	2

II: How Could This Program Be Made to Work Better?
Are resources being used optimally?
Where is there a critical lack of feedback?
What are the barriers to improvement?
What are the critical weaknesses?

Useful Methods	Relevant Tools	General Purpose	References*
Operations Research	Assignment model Transportation model Queueing theory	To maximize the use of existing resources	3, 16, 17
Service Delivery Assessment	Qualitative methods Debriefings Briefings	To provide policymakers near-immediate feedback about the conditions and activities of the program	9, 10
Hearings	Committee hearings Panel reviews	To gather and share testimony from involved and affected parties about the program	24

(Cont.)

Table 1 continued

III: What are the Outcomes of the Program?
What objectives are/are not met?
What "side effects" does the program appear to have?

Useful Methods	Relevant Tools	General Purpose	References*
Achievement Testing	Experimental designs Quasi-experimental designs	To assess if program has statistically significant effect on participants' skills or knowledge	
Survey Research	Questionnaires Surveys Interviews	To assess the perceptions and feelings of individuals and groups about the program	18
Product Evaluation	Critical competitor Systematic check for side effects	To make comparative judgments about product's overall quality	19
Exploratory Data Analysis	Stem and leaf displays Functional transformations Box plots	To make a search for unanticipated results and patterns in accumulated data	2

IV: What Important Variations Are There in the Program's Activities or Effects?
To what extent are different groups affected in different ways?
In what ways has the program varied over time?
How do the program's resources, services, or outcomes vary geographically?

Useful Methods	Relevant Tools	General Purpose	References*
Hearings	Committee hearings Jury methods	To gather testimony from advocates of different points of view	24
Document Analysis	Tracking Legislative history	To trace over time the changes in the operations or character of a program	2, 4, 23
Geographic Methods	Geocode analysis Trend surface analysis Social area analysis	To assess the distribution of program parameters over regional areas	14, 21

VII: How Worthwhile is the Program?
Overall, how good is the program?
Is the program cost-effective?

Useful Methods	Relevant Tools	General Purpose	References*
Product Evaluation	Needs assessment Cost analysis Synthesis procedures	To come to an overall judgment about a program's quality; to aid in decision making	19
Cost Analysis	Feasibility studies Cost-effectiveness analysis Cost-benefit analysis	To render questions of cost into useful forms; to generate information to aid decision making	12, 13, 20

(Cont.)

Table 1 continued

B. Methods to Help in the Design of the Evaluation

I: What Should the Evaluation Focus On?
What are the critical or pay-off issues?
What dimensions are important to include in the study?

Useful Methods	Relevant Tools	General Purpose	References*
Product Evaluation	Checklist	To aid the evaluator in making a comprehensive assessment of a program	19
Investigative Journalism	Quick study	To get the "lay of the land" and review relevant background	8, 15
Case Study	Observation and interview techniques	To gain insight into the whole by studying a single part	
Document Review	Legislative history Content analysis	To learn the historical or legislated intent of a program; to discover program themes or characteristics	2, 4

II: How Can We Move from the General to the Specific?
Do we agree on the meaning of key terms?
How specific do we wish to be?

Useful Methods	Relevant Tools	General Purpose	References*
Philosophical Analysis	Concept analysis	To clarify thinking about general and abstract questions; to see how concepts function in language and thought	1, 7, 11

*See reference list at end of this guide.

Table 2
Tools for the Evaluator

A. Tools for Gathering Information

Tools	Methods	Purpose
Investigative	Investigative Journalism	To confirm hunches; discover new leads
In-depth interview	Case Study	To probe: to gain insight
Testimony	Committee Hearings Panel Reviews	To gather evidence and viewpoints of different interests
Observation	Case Study Phenomenonology SDA	To obtain "snapshops" of reality; to discover patterns
Document review and tracking	Investigative Journalism Legislative History	To substantiate inferences; to learn history of issue or program
Achievement tests	Experimental Design	To determine if groups are statistically different
Operational tests	Product Evaluation	To measure the qualities of performance
Surveys and questionnaires	Market Research	To discover the distribution of opinion
Photographs	Photography	To capture images of reality

(Cont.)

Table 2 continued

B. Tools for Analyzing Information

Tools	Methods	Purpose
Factor analysis t-test	Statistical Analysis	To determine if observed differences are statistically significant
Cost analysis: feasbility effectiveness utility benefit	Cost Analysis	To determine if programs are feasible, or to measure costs against results
Operational analysis: assignment transportation queueing	Operations Research	To determine maximum use of resources; to minimize costs
Box plot Function transformation Stem and leaf display	Exploratory Data Analysis	To discover relationships, and patterns hidden in accumulated data
Geocode, Trend surface, and Social area analyses	Geographic Methods	To portray the spatial distribution of program variables
Thematic matrix analysis	Literary Criticism	To identify predominant themes
Concept analysis	Philosophy	To clarify thinking, language and ideas
Content analysis Tracking	Document Analysis	To substantiate themes; to substantiate a hypothesis
Debriefing	Service Delivery Assessment	To arrive at consensus of perceptions
Connoisseurship	Criticism	To offer personal, expert analysis and opinions
Hearings	Government Committee Hearings	To synthesize evidence in an open public format
Juries	Legal Proceedings	To judge evidence in the form of adversary testimony

(Cont.)

Table 2 continued

C. Tools for Communicating the Findings

Tools	Methods	Purpose
Narrative Prose	Storytelling	To convey the reality, humanness of program
Briefs	Journalism	To convey highlights in headline form
Graphics	Art/Design	To translate information into clear, insightful, graphic form
Maps	Geography	To illustrate relationships using mapping formats
Pictures	Photography	To use pictures to heighten sense of program reality
Oral Briefings	Service Delivery Assessment	To give oral presentation of findings
Hearings	Committee Hearings	To present all testimony and evidence publicly
Vignettes	Case Study	To present in writing typical illustrative scenarios

REFERENCES

1. Caulley, D. N. (1981). Concept analysis in evaluation (ROEP Paper and Report Series No. 61). Portland, OR: Northwest Regional Educational Laboratory.

2. Caulley, D. N. (1982). Legislative history and evaluation, Evaluation and Program Planning, 5, 45-52.

3. Caulley, D. N. (1982). The use of assignment and transportation models in evaluation (ROEP Paper and Report Series No. 68). Portland, OR: Northwest Regional Educational Laboratory.

4. Caulley, D. N. (1983). Document analysis in program evaluation, Evaluation and Program Planning, 6, 19-29.

5. Della-Piana, G. M. (1981). Literary and film criticism. In N. L. Smith (Ed.), Metaphors for evaluation: Sources of new methods. Beverly Hills, CA: Sage Publications.

6. Della-Piana, G. M. (1982). Film criticism and microcomputer courseware evaluation. In N. L. Smith (Ed.), Field assessments of innovative evaluation methods (New Directions in Program Evaluation Series). San Francisco: Jossey-Bass.

7. Gowin, D. B. (1981). Philosophy. In N. L. Smith (Ed.), Metaphors for evaluation: Sources of new methods. Beverly Hills, CA: Sage Publications.

8. Guba, E. G. (1981). Investigative reporting. In N. L. Smith (Ed.), Metaphors for evaluation: Sources of new methods. Beverly Hills, CA: Sage Publications.

9. Hendricks, M. (1981). Service delivery assessment: Qualitative evaluations at the cabinet level. In N. L. Smith, Federal efforts to develop new evaluation methods (New Directions for Program Evaluation Series). San Francisco: Jossey-Bass.

10. Hendricks, M. (1982). Oral policy briefings. In N. L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: Sage Publications.

11. Lane, C. A. (1982). Using the tools of philosophy: Metaphor in action. In N. L. Smith (Ed.), Field assessments of innovative evaluation methods (New Directions in Program Evaluation Series). San Francisco: Jossey-Bass.

12. Levin. H. M. (1981). Cost analysis. In N. L. Smith (Ed.), New techniques for evaluation. Beverly Hills, CA: Sage Publications.

13. Levin, H. M. (1983). Cost-effectiveness: A primer. Beverly Hills, CA: Sage Publications.

14. Monk, J. J. & Hastings, J. T. (1981). Geography. In N. L. Smith (Ed.), Metaphors for evaluation: Sources of new methods. Beverly Hills, CA: Sage Publications.

15. Nelson, D. E. (1982). Investigative journalism methods in educational evaluation. In N. L. Smith (Ed.), Field assessments of innovative evaluation methods (New Directions in Program Evaluation Series). San Francisco: Jossey-Bass.

16. Page, E. B. (1979). Educational evaluation through operations research (ROEP Paper and Report Series No. 30). Portland, OR: Northwest Regional Educational Laboratory.

17. Page, E. B. (1979). Operations research as a metaphor for evaluation (ROEP Paper and Report Series No. 15). Portland, OR: Northwest Regional Educational Laboratory. (ERIC Document Reproduction Service No. ED 206 681)

18. Rasp, A. Jr. (1982). Interviewing to augment large scale survey data: The Washington high school and beyond story (ROEP Paper and Report Series No. 71). Portland, OR: Northwest Regional Educational Laboratory.

19. Scriven, M. (1981). Product evaluation. In N. L. Smith (Ed.), New techniques for evaluation. Beverly Hills, CA: Sage Publications.

20. Smith, J. K. (1983). Case reports of Northwest Regional Educational Laboratory cost studies (ROEP Paper and Report Series No. 82). Portland, OR: Northwest Regional Educational Laboratory.

21. Smith, N. L. (1979). Techniques for the analysis of geographic data in evaluation. Evaluation and Program Planning, 2, 119-126.

22. Smith, N. L. (1980). Bibliography of evaluation utilization (ROEP Paper and Report Series No. 39). Portland, OR: Northwest Regional Educational Laboratory.

23. Smith, N. L. (1982). Investigative tracking in library evaluation. In N. L. Smith (Ed.), Field assessments of innovative evaluation methods (New Directions in Program Evaluation Series). San Francisco: Jossey-Bass.

24. Stenzel, N. (1979). Committee hearings as an evaluation format. In N. L. Smith (Ed.), Field assessments of innovative evaluation methods (New Directions in Program Evaluation Series). San Francisco: Jossey-Bass.

25. Templin, P. A. (1982). Still photography in evaluation. In N. L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: Sage Publications.

26. Wachtman, E. L. (1982). Storytelling: The narrative structure of educational evaluation. In N. L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: Sage Publications.

DESIGN STRATEGY FOR INCORPORATING NEW METHODS

A more extensive and complex piece of art (e.g., a symphony) cannot be created by starting at the beginning and going on in a linear way to the end. On the contrary, it is necessary to continuously keep the totality in mind, and go from the whole to the parts and from the parts back to the whole.

Norberg Shulz (Intensions in Architecture)

THE DESIGN PROBLEM

When you set out to evaluate an educational program, how do you know what to do? How do you decide what kinds of activities your study will include? What approaches or methods do you consider using?

One way of deciding what to do is to identify yourself with a method or an approach ("I am an expert in experimental designs") or ("I do case study research"). This approach certainly resolves many methodological questions and works fine, as long as you are careful to study only questions that are amenable to the methods you specialize in. However, the lessons of recent evaluation history appear to point out that one-dimensional evaluations, even when they are done well, are often of very limited value. Most programs are multi-dimensional and require a correspondingly multifaceted understanding.

A second way of deciding what to do when you find yourself in the position of the evaluator is to settle on an overall purpose, or a set of goals, for the evaluation. This purpose can then guide the choice of activities and methods. For instance, an evaluator can adopt the purpose of measuring the extent to which the program's objectives have been achieved. While this purpose is, in fact, commonly chosen, it is clearly not unique. A slightly different goal (often adopted by state evaluation agencies) is to evaluate the extent to which the originally legislated intent of a program has been realized.

There are other evaluation goals worth considering, as well. You might, as in an accreditation study, wish to certify a program's quality by comparing it to some external standards. You might wish to see to what extent a program is fulfilling students' actual needs, or you might decide that it is most important to discover ways to improve the program. It is clear that there are many possible goals for an evaluation--many of them quite worthwhile, and each one requiring a different set of methods and activities. Where do we find guidelines for choosing among them?

A third way of deciding what to do is to consider the audience of the evaluation. By considering the needs of decision makers, program developers, staff, students, interested researchers, and other affected special interest groups, the evaluator may choose method(s) which enhance the impact of the evaluation.

Another source of guidance for the evaluator is the Program itself, since its nature is clearly going to have an effect on the kinds of methods the evaluator chooses. For example, an in-depth case study may be appropriate for the evaluation of a controversial graduate theological program, but not as useful in the study of the utility of a high school typing course.

How does one then decide what to do? All of the above factors--the skills of the evaluator, the purpose of the evaluation, the nature of the program, and the needs of the audiences--influence the approach and methods the evaluator may select.

The evaluator typically faces a situation in which many factors influence and make demands upon the evaluation design. That is, neither the evaluation task nor the program being evaluated are well defined. This is exactly the kind of problem that designers face as they try to create functional forms that satisfy many different criteria. It is the general nature of the design problem that makes the architect a useful metaphor for the evaluator.

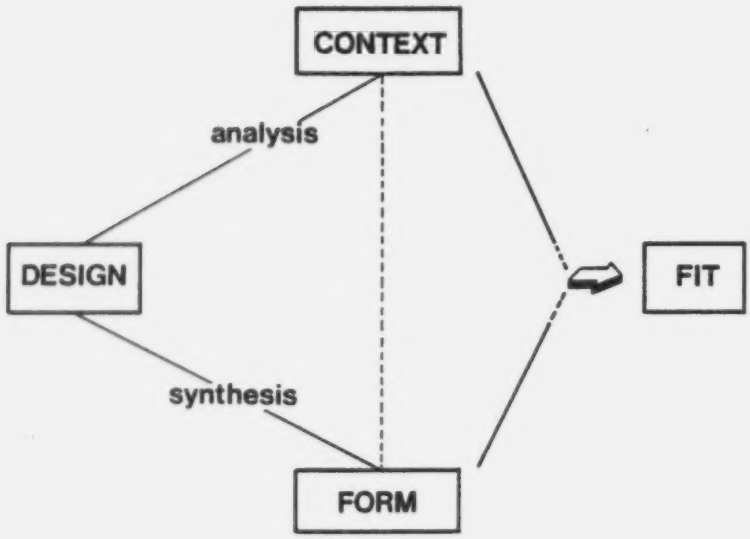
In the following section, the process that an architect goes through in defining and solving a design problem is described in terms of a general model, which is shown to be analogous to the evaluation process.

THE ARCHITECT

The architect is concerned with the process of design--the process of inventing physical things that display order and organization in forms that respond to functional needs. Successful design requires that the architect operate both in a linear, analytical fashion and in a more intuitive, holistic mode. The architect must possess a wide range of technical skills and use the methods of many varied disciplines.

In moving from the initial presentation of the problem to the development of the final design form, the architect goes through a process which can be reconstructed in a general way, as shown in Figure 1.

Figure 1



Context

The context is the name given to the multidimensional problem space that the architect faces. The context is the set of all factors which make demands upon the final design form. The context includes physical, legal, psychological, and philosophical factors. All together these factors define the boundaries of possible solutions. For example, the context for the architect includes such factors as the client's desires, the functions the final form must serve, the resources available (including funds, material, personnel, time . . .), the nature of the site (including soil type, slope, exposure, access . . .), and any social or legal constraints (such as building codes, zoning laws, neighborhood tradition . . .). All of these factors are elements of the architect's problem context. While they influence the direction and shape of the architect's solution, they do not uniquely define it.

Design

In the first phase of the design process, the architect carries out an analysis of the context. Information is gathered and separated into components along different dimensions. For example, the architect may examine the "site dimensions"

collecting information about the soil, the topography, the vegetation, the orientation to the sun and the wind, the variation with season, and so forth. Studying the context in an open-ended way, the architect discovers which dimensions are important, probes into the component factors of that dimension, and even begins to learn about how different components interact. (For example, in deciding where to place a house on the crest of an exposed hill, the architect may learn about the trade-offs between energy efficiency and the quality of the view.)

In the second phase of the design, the architect begins to formulate outlines and pieces of the design. Priorities emerge and vague goals are stated (such as "the design should create an ambiance of excellence and tradition"). The philosophy, the style, and the basic elements of the design begin to take shape in the architect's mind.

The Form

The form is the realization of the design process; it is a solution which it is hoped fulfills those dimensions of the need most demanding of attention. The form is arrived at through a process of synthesis in which the ideas that arose in the design process through analysis are combined to form a coherent whole. Since the demands or factors influencing the form interact with each other, and since the final form must simultaneously satisfy as many of these demands as possible, a kind of gestalt or integrated vision of the final design is required. In achieving this integration, the architect may begin by sketching important conceptual relationships, by then moving to two-dimensional diagrams, and finally by producing three-dimensional models.

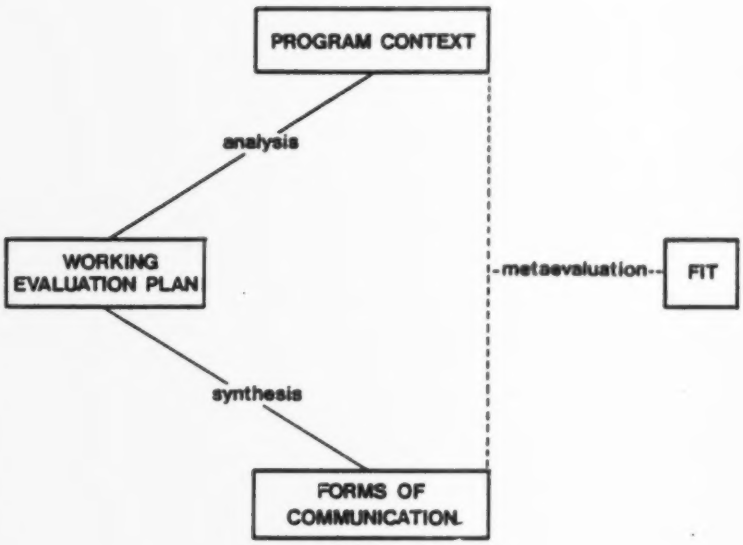
The Fit

The fit is the degree of congruence between the form and the context. It is a measure of acceptability. Out of the evaluation of the fit, judgments are made about how well the design has met the many demands of the context. Standards of consistency, integrity, and aesthetics are applied. Poor designs lead to such judgments as "arbitrary, obsolete, incongruous, or dysfunctional." Successful designs are judged to be "efficient, compatible, energy-saving, and proportional." Such judgments are based on the contextual factors deemed most important in a given case.

THE EVALUATOR

Very much like the architect, the evaluator faces a problem of finding forms that best fit a partially defined set of contextual demands. The process the evaluator goes through can be viewed as very similar to that of the architect, as illustrated by Figure 2.

Figure 2



Program Context

The context for the evaluator consists of a great variety of factors, including, for example, the following concerns:

The needs of:

- the client or sponsoring agency
- the students or other affected audience
- the teachers or program staff
- other indirectly affected audiences

The nature of:

- the program's goals or intent
- the program's activities
- the educational or institutional setting
- the political setting

The constraints of:

- time (the program's and the evaluator's)
- resources (including money and personnel)
- access
- social or legal requirements
- the evaluator's own evaluation skills
- the evaluation mandate

All of these factors push and pull the evaluator in the attempt to create a balanced and effective evaluation design.

Working Evaluation Plan

As the design is to the architect, so is the working evaluation plan to the evaluator. Like the architect, the evaluator's design comes out of an analysis of the context--that is, out of an insightful understanding of the total evaluation setting. The evaluator's working plan corresponds to the architect's sketches, and yet the working plan is a more dynamic entity. As the evaluation proceeds through the design phase and into its implementation, more is learned about the context. Important factors that were unknown will emerge, and what was thought to be important may turn out not to be. Accordingly, the working plan is never fixed, but rather is continuously evolving and changing.

Forms of Communication

The form, the product of the evaluator's design, is not really physical. Even the "final report" is more a means to an end than an end in itself. What the evaluator ultimately wishes to do is to communicate--to inform, educate, inspire, arouse, or otherwise produce a beneficial impact upon the appropriate people. Thus, it is effective interaction that is the product, the final form, of a successful evaluation design.

Fit

The fit is the degree to which the evaluator is able to understand and meet the demands of the situation. By constantly monitoring what is happening during the evaluation (a kind of dynamic self meta-evaluation), the evaluator can continually update the working plan and flexibly adjust the design to respond to what is needed.

SUMMARY
A Few General Principles

When evaluation is treated as a design problem, the vision of the evaluator's task is expanded, and the evaluator needs to become more creative, flexible, and artistic. There is less certainty, and rather than have a set of algorithms to follow, the best the evaluator can hope for are some general guidelines to go by. As discussed earlier, out of the general nature of the design process we can, with a little help from actual experience, extract the following general principles of "ideal" evaluation design:

1. Begin by <u>not</u> knowing what to do.

   This takes an act of courage on the evaluator's part. Many times there are strong pressures (both internal and external) to restrict the evaluation to a predetermined task or method. Often the evaluator will need to educate and convince clients that, to meet their ultimate needs, the evaluation will have to be different from what they initially envisioned. Even if, ultimately, the evaluator does exactly follow a prescribed course, an open-minded look at the beginning can only increase the overall quality of the work that follows.

2. Begin with what <u>is</u> happening.

   This is really a corollary to the first principle. Often the evaluator enters the scene blinded by someone else's concerns, the description of the program and its goals, or self-imposed ideas about what is and is not happening. All of these are legitimate concerns and questions--they should ultimately be addressed--but they are not necessarily good starting points. By beginning with what is happening, by studying the program as it is, the evaluator becomes grounded in a personal knowledge of the reality of the program. Then, through observation, discussion, document analysis, and serendipity, the evaluator can discover what in the setting (context) is and is not important.

3. Work from the general to the specific.

   Issues, questions, and important dimensions (as well as ways to proceed) may at first be only vaguely sensed. As in the design of a building, the final design will be more complete and successful if the evaluator's vision is allowed to mature slowly and change along the way.

4. Use both analytic and integrative modes of thinking.

   Analysis is needed to discover and separate out the important dimensions of the scene, while an integrated vision is required to create an encompassing vision of the evaluation in its entirety. And these two modes of thought interact. When analysis yields a new factor to consider, the overall vision of the evaluation may change. When the overall vision changes, new directions for analysis may be indicated.

5. Formulate general questions to guide the evaluation.

   Out of the analysis of the context, the evaluator begins to sense what is and what is not worth addressing. This sense can be made more specific by formulating a few general questions to guide the activities and methods of the evaluation. Questions such as "How can we understand what is happening in the program?" or "What are the barriers to this program's effectiveness?" are at about the right level of detail.

6. Use evaluation methods flexibly.

   The architect must draw upon and skillfully use methods from a wide range of disciplines (art, mechanical drawing, engineering, sociology . . .). All of the techniques of these disciplines are needed and used by the architect. Guided by what needs to be done, a skillful architect even takes delight in the creative use of many different skills and methods. The evaluator would do well to emulate the architect in terms of the flexible use of methods.

7. Monitor, recycle, and rethink.

   For some reason, evaluation plans seem to be cast in bronze early on. Even if the charted course makes no sense to anyone, the evaluation often proceeds as scheduled. However, if the evaluator begins with the idea of working with a sketch--a plan that is by design going to evolve and change--then the final evaluation is more likely to fit with what is needed.

REFERENCE

Ferguson, G. (1981). Architecture. In N. L. Smith (Ed.), Metaphors for evaluation. Beverly Hills, CA: Sage Publications.